Some interesting dual code properties of convolutional encoder for standards self recognition

Mélanie Marazin[1-2] Roland Gautier[1-2] Gilles Burel[1-2]
[1]Université Européenne de Bretagne, France.
[2]Université de Brest; CNRS, UMR 3192 Lab-STICC, ISSTB, 6 avenue Victor Le Gorgeu, CS 93837, 29238 Brest cedex 3, France

***Abstract*—For enhancement of the quality of digital transmissions, standards are in continual evolution, which generates compatibility problems. Cognitive radio systems permit one to solve this problem through the design of intelligent receivers. However, such receivers must be able to adapt themselves to a specific transmission context. This requires the development of new methods in order to blindly estimate error-correcting codes. Coding schemes like turbocode, composed of convolutional codes, belong to a family of error-correcting codes in use in many standards. In most of the methods designed to identify convolutional encoders the algebraic properties are used implicitly. But usually, these dedicated properties are neither explicated, nor detailed, nor demonstrated. The study reported here investigates the algebraic properties of convolutional encoders, useful for blind recognition methods in the cognitive radio context and more specially the algebraic relationships between different forms of a convolutional code and its corresponding dual code. Moreover, some simulation results are presented to show the relevance of these properties for the blind identification of the convolutional encoder.**

I. Introduction

In order to meet the expectations and transmission constraints about data rate or readability generated by new applications, digital communication systems are in constant evolution. With the fast development of new communication standards, the design of intelligent receivers has become a must. Indeed, such receivers can adapt to a specific transmission context through a blind estimate of the transmitter parameters. This requires the development of new methods to blindly estimate error-correcting codes known to enhance the quality of communications by enabling the binary data stream to better withstand channel impairments such as a noisy transmission channel, interferences or channel fading. For this purpose, they introduce some redundancy in the informative binary data stream.

A literature review shows that most of the methods dedicated to the blind identification of convolutional encoder use both some algebraic properties of convolutional encoders and those of their dual codes. These considerations led us to study certain algebraic properties of convolutional encoders used in blind recovery methods. The paper is organized as follows: section II introduces some properties of convolutional encoders as well as the notion of equivalent encoder. Then, section III investigates the relationship between the dual code and the code. Finally, section IV shows the interest of these algebraic properties on the blind recognition methods of convolutional encoders. Our conclusion is drawn in section V.

Email: {melanie.marazin, roland.gautier, gilles.burel}@univ-brest.fr

II. Convolutional encoder

A convolutional encoder is defined by a set of three parameters, respectively denoted by n, k and K, where n is the number of outputs, k is the number of inputs and K is the constraint length, plus a $(k \times n)$ generator matrix [denoted as $G(D)$] such that

$$G(D) = \begin{bmatrix} g_{1,1}(D) & \cdots & g_{1,n}(D) \\ \vdots & \cdots & \vdots \\ g_{k,1}(D) & \cdots & g_{k,n}(D) \end{bmatrix} \quad (1)$$

where $g_{i,j}(D)$ are generator polynomials or generator rational functions.

Let us denote by $m(D)$ and $c(D)$ the input and output sequences, respectively. So, the relation between them is expressed as

$$c(D) = m(D).G(D). \quad (2)$$

A. Equivalent encoder

One of the most important properties in the error correction theory is the notion of equivalent encoder. Indeed, a given convolutional code can be encoded by several different encoders. Moreover, it has both systematic rational generator matrices (where the entries are rational functions) and polynomial generator matrices (where all the entries are polynomials). The systematic rational generator matrices give encoders with 'feedback'. Let us denote the encoders with feedback by RSC (Recursive and Systematic Code) and those with no feedback by NRNSC (Non-Recursive and Non-Systematic Code). Properties of equivalent encoders were given in earlier papers [1], [2].

***Definition** 1: Two convolutional generator matrices $G(D)$ and $G'(D)$ are equivalent if they encode the same code, C. Two convolutional encoders are equivalent if their generator matrices are equivalent.*

***Theorem** 1: Two rate $r = k/n$ code generator matrices $G(D)$ and $G'(D)$ are equivalent if, and only if there is a $k \times k$ nonsingular matrix $T(D)$ such that*

$$G(D) = T(D).G'(D). \quad (3)$$

***Theorem** 2: Every convolutional generator matrix is equivalent to a systematic rational encoding matrix.*

For a convolutional encoder, the choice of generator matrix is essential. Indeed, the various generator matrices give different complexities both in encoding and decoding procedures. Among the generator matrices, the class corresponding to

those termed as catastrophic ones must be avoided. Indeed, with such a catastrophic matrix, a small number of channel errors may generate an unlimited number of errors after decoding. Furthermore, within a class of generator matrices for a code, the most used encoder is the one with the most desirable structural properties. These generator matrices, which describe optimal convolutional encoders, have good algebraic properties ([1], [3]) that can be judiciously exploited for blind identification.

B. Relation between the NRNSC and the RSC encoders

Evidence of the relationship between the generator matrix of an NRNSC encoder and that of its RSC equivalent encoder was given in an earlier paper [4]. It is briefly recalled hereafter prior to the study of the relationship between the dual code and the code. For this, let us denote by $G_{NRNSC}(D)$ a generator matrix of an NRNSC encoder and by $G_{RSC}(D)$ a generator matrix of its RSC encoder. Then, the matrix $G_{NRNSC}(D)$ is defined by

$$G_{NRNSC}(D) = \begin{bmatrix} g_{1,1}(D) & \cdots & g_{1,n}(D) \\ \vdots & \cdots & \vdots \\ g_{k,1}(D) & \cdots & g_{k,n}(D) \end{bmatrix}. \quad (4)$$

On condition to denote by $L(D)$ a $(k \times k)$ sub-matrix of $G_{NRNSC}(D)$, such that

$$L(D) = \begin{bmatrix} g_{1,1}(D) & \cdots & g_{1,k}(D) \\ \vdots & \cdots & \vdots \\ g_{k,1}(D) & \cdots & g_{k,k}(D) \end{bmatrix} \quad (5)$$

the RSC equivalent matrix is

$$G_{RSC}(D) = \frac{1}{\det L(D)}.\text{adj } L(D).G_{NRNSC}(D) \quad (6)$$

where $\det L(D)$ is the determinant of $L(D)$, and adj $L(D)$ is the adjoint matrix of $L(D)$. The matrix, $G_{RSC}(D)$, is a $(k \times n)$ matrix such that

$$G_{RSC}(D) = \begin{bmatrix} 1 & & \frac{f_{1,k+1}(D)}{f_{1,1}(D)} & \cdots & \frac{f_{1,n}(D)}{f_{1,1}(D)} \\ & \ddots & \vdots & \cdots & \vdots \\ & 1 & \frac{f_{k,k+1}(D)}{f_{1,1}(D)} & \cdots & \frac{f_{k,n}(D)}{f_{1,1}(D)} \end{bmatrix} \quad (7)$$

where $f_{i,j}(D)$ and $f_{1,1}(D)$ are termed, respectively, as the generator polynomials of $G_{RSC}(D)$, $\forall i = 1, \cdots, k$ and $\forall j = k+1, \cdots, n$, and the feedback polynomial.

According to (6), the $f_{i,j}(D)$ polynomials are obtained by multiplying the ith row of adj $L(D)$ with the jth column of $G_{NRNSC}(D)$. An adjoint matrix is defined by

$$\text{adj } L(D) = \begin{bmatrix} Cof_{1,1}(D) & \cdots & Cof_{k,1}(D) \\ \vdots & \cdots & \vdots \\ Cof_{1,k}(D) & \cdots & Cof_{k,k}(D) \end{bmatrix} \quad (8)$$

where $Cof_{p,i}(D)$ is a determinant of a sub-matrix of $L(D)$ obtained by deleting the pth row and the ith column of $L(D)$. Each generator polynomial $f_{i,j}(D)$ is such that

$$f_{i,j}(D) = \sum_{p=1}^{k} g_{p,j}(D).Cof_{p,i}(D). \quad (9)$$

Thus, $f_{i,j}(D)$ corresponds to the determinant of a sub-matrix of $L(D)$ obtained by deleting the ith row and the ith column. These generator polynomials are a k-order minor of $G_{NRNSC}(D)$ matrix.

Most methods of blind identification allow one to find the NRNSC equivalent form of the really used encoder. Thus, the relation between a generator matrix of an NRNSC encoder and its RSC equivalent encoder expressed in (6) is paramount in blind recovery. Indeed, if the encoder is in the RSC form, this relation permits one to get the true generator matrix.

III. Relationship between the dual code and the RSC encoder

In this section, focus is on the relation between the generator matrix of a dual code and the generator matrix of the RSC encoder.

A. Dual code

A convolutional encoder can also be described by a dual code generator matrix termed parity check matrix, whose properties ([2], [5]) are known as excellent in the blind identification of convolutional encoders.

***Property** 1: Let $G(D)$ be a generator matrix of C. If an $((n-k) \times n)$ polynomial matrix, $H(D)$, is a parity check matrix of C, then*

$$G(D).H^T(D) = 0. \quad (10)$$

***Corollary** 1: Let $H(D)$ be a parity check matrix of C. The output sequence $c(D)$ is a codeword sequence of C if and only if*

$$c(D).H^T(D) = 0. \quad (11)$$

The methods dedicated to the blind estimation of convolutional encoders are based on the algebraic properties of convolutional codes ([6], [7]). In such blind recovery methods, the first step is the identification of the code parameters (k, n and K). It is followed by the identification of the code parity check matrix. At last, a generator matrix of its NRNSC encoder can be deduced from this parity check matrix. But, as observed for convolutional encoders, the dual code is described by many matrices. In blind recovery, getting a parity check matrix with good algebraic properties is essential to deduce a generator matrix of its NRNSC encoder. This requirement drove us to gain more insight into the parity check matrix, $H(D)$, used to identify a convolutional encoder.

This parity check matrix is an $((n-k) \times n)$ matrix such that

$$H(D) = \begin{bmatrix} h_{1,1}(D) & \cdots & h_{1,k}(D) & h_0(D) & & \\ \vdots & \cdots & \vdots & & \ddots & \\ h_{n-k,1}(D) & \cdots & h_{n-k,k}(D) & & & h_0(D) \end{bmatrix} \quad (12)$$

where $h_0(D)$ and $h_{i,j}(D)$, $\forall i = 1, \cdots, n-k$ and $\forall j = 1, \cdots, k$, are the generator polynomials of $H(D)$.

B. Relation between $H(D)$ and $G_{RSC}(D)$

Let us consider the previous $H(D)$ matrix (12) in the case where it is composed of the generator polynomials of an RSC encoder so that

$$H(D) = \begin{bmatrix} f_{1,k+1}(D) & \cdots & f_{k,k+1}(D) & f_{1,1}(D) & & \\ \vdots & \cdots & \vdots & & \ddots & \\ f_{1,n}(D) & \cdots & f_{k,n}(D) & & & f_{1,1}(D) \end{bmatrix}. \tag{13}$$

To show that this matrix (13) is a parity check matrix of the NRNSC encoder equivalent to the previous RSC encoder, let us denote by $G(D)$ an NRNSC generator matrix and by $f_{i,j}(D)$ the generator polynomials of its RSC equivalent encoder.

Let us denote by $R(D)$ a $(k \times (n-k))$ matrix defined by

$$\begin{aligned} R(D) &= G(D).H^T(D) \\ &= \begin{bmatrix} r_{1,1}(D) & \cdots & r_{1,n-k}(D) \\ \vdots & \cdots & \vdots \\ r_{k,1}(D) & \cdots & r_{k,n-k}(D) \end{bmatrix} \end{aligned} \tag{14}$$

where the polynomials, $r_{i,m}(D)$, $\forall i = 1, \cdots, k$ and $\forall m = 1, \cdots, n-k$, are such that

$$r_{i,m}(D) = \sum_{j=1}^{k} g_{i,j}(D).f_{j,k+m}(D) + g_{i,k+m}(D).f_{1,1}(D). \tag{15}$$

According to (9), the polynomials $r_{i,m}(D)$ are

$$\begin{aligned} r_{i,m}(D) = &\sum_{j=1}^{k}\sum_{p=1}^{k} g_{i,j}(D).g_{p,k+m}(D).Cof_{p,j}(D) \\ &+ \sum_{p=1}^{k} g_{i,k+m}(D).g_{p,1}(D).Cof_{p,1}(D) \end{aligned}. \tag{16}$$

Thus, splitting $r_{i,m}(D)$ into two polynomials leads to

$$r^1_{i,m}(D) = \sum_{j=1}^{k}\sum_{p=1}^{k} g_{i,j}(D).g_{p,k+m}(D).Cof_{p,j}(D) \tag{17}$$

and

$$r^2_{i,m}(D) = \sum_{p=1}^{k} g_{i,k+m}(D).g_{p,1}(D).Cof_{p,1}(D) \tag{18}$$

- Study of the polynomial $r^1_{i,m}(D)$

$$r^1_{i,m}(D) = \sum_{p=1}^{k}\sum_{j=1}^{k} g_{i,j}(D).Cof_{p,j}(D).g_{p,k+m}(D). \tag{19}$$

For a fixed value of p, the part $\sum_{j=1}^{k} g_{i,j}(D).Cof_{p,j}(D)$ corresponds to the determinant

$$\det \begin{pmatrix} g_{1,1}(D) & \cdots & g_{1,k}(D) \\ \vdots & \cdots & \vdots \\ g_{p-1,1}(D) & \cdots & g_{p-1,k}(D) \\ g_{p+1,1}(D) & \cdots & g_{p+1,k}(D) \\ \vdots & \cdots & \vdots \\ g_{k,1}(D) & \cdots & g_{k,k}(D) \\ g_{i,1}(D) & \cdots & g_{i,k}(D) \end{pmatrix}. \tag{20}$$

This determinant has two different values

- If $p \neq i$

$$\sum_{j=1}^{k} g_{i,j}(D).Cof_{p,j}(D) = 0 \tag{21}$$

- If $p = i$

$$\sum_{j=1}^{k} g_{i,j}(D).Cof_{p,j}(D) = \det L(D) \tag{22}$$

By resuming (21) and (22), the polynomial $r^1_{i,m}(D)$, (17), is such that

$$\begin{aligned} r^1_{i,m}(D) &= \sum_{j=1}^{k}\sum_{p=1}^{k} g_{i,j}(D).g_{p,k+m}(D).Cof_{p,j}(D) \\ &= \det L(D).g_{i,k+m}(D) \end{aligned}. \tag{23}$$

- Study of the polynomial $r^2_{i,m}(D)$

$$\begin{aligned} r^2_{i,m}(D) &= \sum_{p=1}^{k} g_{i,k+m}(D).g_{p,1}(D).Cof_{p,1}(D) \\ &= \det L(D).g_{i,k+m}(D) \end{aligned}. \tag{24}$$

According to (23) and (24), the polynomial $r_{i,m}(D)$, (16), is such that

$$r_{i,m}(D) = 0 \quad \forall i, m \tag{25}$$

Thus, the matrix $R(D)$ is composed of zero polynomials. Consequently, the matrix defined in (12), $H(D)$, is a parity check matrix of the encoder. In practice, it is usual to only employ optimal convolutional encoders because their error correction capabilities are the highest. In the blind recovery context, the algebraic properties of these optimal convolutional encoders can be judiciously exploited. In fact, this generates strong properties on a generator matrix of an NRNSC or RSC encoder. In this section, we proved that the generator polynomials of the encoder RSC correspond to the polynomials of the parity check matrix of its equivalent NRNSC encoder. Thus, the parity check matrix, H(D), used to deduce the generator matrix of an NRNSC encoder has also excellent properties for blind identification methods.

IV. Simulation results

In [7], an iterative process dedicated to the blind identification of a rate $(n-1)/n$ convolutional encoder in a noisy environment is explained. The principle of this method is to first identify the number of outputs n. Then, a basis of the dual code can be estimated. And finally, the knowledge of these parameters allows to identify a generator matrix. Let us recall the principle of this algorithm.

The first step is to reshape columnwise the received data bit stream under matrix form of size $(M \times l)$, denoted R_l. This matrix is computed for different values of l ($\forall l = 1, \cdots, M/2$) and for each matrix the Gauss Jordan Elimination Trough Pivoting is applied to obtain a lower triangular matrix noted G_l

$$A_l.R_l.B_l = G_l \tag{26}$$

In (26), A_l is an $(M \times M)$ rows permutation matrix and B_l an $(l \times l)$ matrix describing the columns combination. To detect

the value of n, the principle is to find matrices R_l which exhibit a degenerated rank. So, the gap between two matrices R_l which have a dependent columns detected corresponds to n. Then a dual code basis can be built from the matrix B_l and finally with (10) a linear system can be solved to estimate a generator matrix. But, to obtain the generator matrix of the encoder used at the transmitter, it is important to have beforehand identified the parity check matrix in the same form of (13).

Here, the relevance of the algebraic properties of convolutional codes and dual codes in the blind identification methods is studied. An example of a $C(2,1,7)$ convolutional code is taken. This encoder is used in many standards and it is described by the generator matrix and the parity check matrix such that

$$G = (133 \quad 171) \text{ and } H = (171 \quad 133) \tag{27}$$

where polynomials are represented in octal.

To analyse the impact of the true identification of parity check matrix upon the global performances of the blind identification method, two probabilities were defined as follows:

- probability of identifying the true encoder (parameters and generator matrix) denoted $P_{det} - >$ encoder;
- probability of identifying the true parity check matrix denoted $P_{det} - > H$.

Moreover, to evaluate the relevance of the probability of identifying the true encoder obtained, the different probabilities of detection are compared to the code correction capability. For that, let us denote by BER_r the theoretical residual bit error rate obtained after decoding of the corrupted data stream with a hard decision, [2]. In [7], the BER_r is considered as acceptable if it is close to 10^{-5}, since after this limit, the decoded data stream is not clean enough to meet requirements of standard applications.

For the $C(2,1,7)$ code, Fig. 1 shows the different probabilities compared with the channel error probability, denoted P_e, and the limit of the 10^{-5} acceptable BER_r. One should note that our blind identification approach based on the dual code properties of convolutional encoders is pertinent and very impressive regarding to the probability of identifying the true parity check matrix which is greater than 0.95 for a $BER_r < 10^{-5}$ corresponding to $P_e < 2.10^{-2}$, after only one iteration of the iterative process. We can also note that, even if the true parity check matrix is identified, the true encoder can be misidentified, which is clearly visible in Fig. 1 where ($P_{det} - >$ encoder) is always lower than ($P_{det} - > H$) for a $BER_r > 10^{-5}$. Indeed, it is important to identify the true parity check matrix in order to have the possibility to estimate the true encoder. Moreover, an interesting aspect of our approach is the robustness against misestimation of n on the dual code identification and therefore of the true parity check matrix taking into account the properties of convolutional codes. In this example, after having estimated n and H, admisible k and K are tested regarding the properties of convolutional codes and their dual codes. If n is misidentified this can leads to a misidentification of the code, even if the true parity check matrix has been well identified. But this problem can be corrected during the iterative process to improve the probability of identifying the true encoder. Indeed, it is important to noted that the probabilities can be improved by the iterative process of the blind identification method. For example, after five iterations, ($P_{det} - >$ encoder) = 0.97.

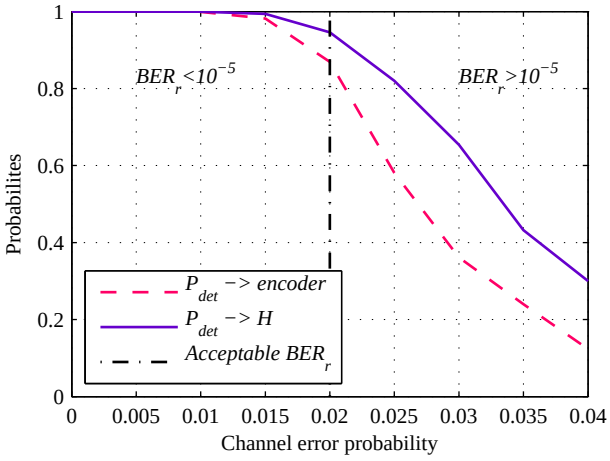

Figure 1. Probabilities of detection compared with P_e

V. Conclusion

This study described some algebraic properties of convolutional encoders and those of their dual codes. It also presented the relation between the NRNSC encoder and the RSC equivalent encoder, and explained the relation between these two generator matrices (RSC and NRNSC) and their parity check matrix. Finally, an analysis of the impact of these algebraic properties upon blind identification methods of convolutional encoders is proposed. This study shows that these algebraic properties are essential for the implementation of blind identification methods.

Acknowledgments

This study was supported by the Brittany Region (France).

References

[1] Forney G. D.: 'Convolutional codes I: Algebraic structure', IEEE Transactions on Information Theory, 1970, 16, (6), pp. 720-738

[2] Johannesson R., and Zigangirov K.Sh.: 'Fundamentals of Convolutional Coding' (IEEE Press, 1999)

[3] McEliece R.J.: 'The algebraic theory of convolutional codes', in Handbook of coding theory, S. (Ed.): 'V.S. Pless and W.C. Huffman' (Elsevier, 1998), pp. 1065-1138

[4] Marazin M., Gautier R., and Burel G.: 'Blind recovery of the second convolutional encoder of a turbo-code when its systematic outputs are punctured', MTA Review, 2009, XIX, (2), pp. 213-232

[5] Forney G. D.: 'Structural Analysis of Convolutional codes via Dual Codes', IEEE Transactions on Information Theory, 1973, 19, (4), pp. 512-518

[6] Barbier J., Sicot G., and Houcke S.: 'Algebraic approach for the reconstruction of linear and convolutional error correcting codes', International journal of applied mathematics and computer sciences, 2006, 2, (3), pp. 113-118

[7] Marazin M., Gautier R., and Burel G.: 'Dual code method for blind identification of convolutional encoder for cognitive radio receiver design', Proc. in the 5^{th} IEEE Broadband Wireless Access Workshop, IEEE GLOBECOM, Honolulu, Hawaii, USA, November 2009, pp. 1-6.